



# The role of age-mixing patterns in HIV transmission dynamics: Novel hypotheses from a field study in Cape Town, South Africa

Roxanne Beauclair<sup>a,b,\*</sup>, Niel Hens<sup>c,d</sup>, Wim Delva<sup>a,b,c,e,f</sup>

<sup>a</sup> International Centre for Reproductive Health, Ghent University, Gent, Belgium

<sup>b</sup> The South African Department of Science and Technology-National Research Foundation (DST-NRF) Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa

<sup>c</sup> Center for Statistics, I-BioStat, Hasselt University, Martelarenlaan 42, BE3500 Hasselt, Belgium

<sup>d</sup> Centre for Health Economics Research and Modelling Infectious Diseases and Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute, University of Antwerp, Antwerp, Belgium

<sup>e</sup> Rega Institute for Medical Research, KU Leuven, Herestraat 49, 3000 Leuven, Belgium

<sup>f</sup> Department of Global Health Faculty of Medicine and Health Sciences, Stellenbosch University, Stellenbosch, South Africa

## ARTICLE INFO

### Keywords:

Age-disparate relationships  
Age-mixing  
HIV infection  
South Africa  
Sexual behaviour

## ABSTRACT

**Background:** Age-disparate relationships are thought to put young women at increased risk of HIV, though current evidence is inconclusive. Studying population-level age-mixing patterns as well as individual-level measures of age difference variation may provide insight into the persistence and magnitude of the epidemic in South Africa.

**Methods:** We used data from a survey in Cape Town ( $n = 506$ ) to describe age-mixing dynamics in the four population strata of HIV negative and HIV positive male and female participants. Mixed-effects models were used to calculate the average increase in partner age for each year increase in age of participant, the average partner age for 15 year olds, and the between-subject and the within-subject standard deviation of partner ages. We conducted 2000 bootstrap replications of the models. Using negative binomial models, we also explored whether HIV status was associated with participants having a larger range in partner ages.

**Results:** HIV positive women had large variability in partner ages at the population level, and at the individual level had nearly three times the expected range of partner ages compared to HIV negative women. This pattern may increase the potential for HIV transmission across birth cohorts and may partially explain the persistence of the epidemic in South Africa. Young men, who have been previously absent from the age-disparity discourse, also choose older partners who may be putting them at increased risk of HIV infection due to the high HIV prevalence among older age categories of women.

## 1. Introduction

Sexual relationships characterized by men being older than their female partners are thought to play a pivotal role in the magnitude and persistence of the HIV epidemic in Sub-Saharan Africa (Anderson et al., 1992; Chapman et al., 2010; Hallett et al., 2007). These relationships, commonly referred to as age-disparate relationships, may put young women at increased risk of acquiring HIV because the relationships also tend to be marked by power imbalances which may result in inconsistent condom use (Gregson et al., 2002; Maughan-brown et al., 2016; Morrison-Beedy et al., 2013; Ritchwood et al., 2016), higher sex frequencies per unit of time (Beauclair et al., 2016; Gregson et al., 2002; Morrison-Beedy et al., 2013; Ritchwood et al., 2016), and multiple partnership concurrency (Gregson et al., 2002; Maughan-brown et al.,

2016, 2014; Ritchwood et al., 2016; Street et al., 2015). Additionally, in South Africa it has been noted that HIV prevalence for men peaks in older age groups compared to women (Shisana et al., 2014). This implies that if a younger woman chooses partners from an older age group, she will be more likely to select a partner that is HIV positive compared to if she was choosing a partner from a similar age (Gouws et al., 2008; Pettifor et al., 2005).

Despite great potential for age-disparate relationships to increase HIV transmission, some recent high-profile studies directly studying the link between age differences – in “primary” or “most recent” relationships – and HIV incidence did not find evidence of an association (Balkus et al., 2015; Harling et al., 2014; Street et al., 2015). Though, one study in Zimbabwe did find that increasing partner age differences were associated with increased hazards of acquiring HIV (Schaefer

\* Corresponding author at: International Centre for Reproductive Health, Ghent University, Gent, Belgium.  
E-mail address: [Roxanne.Beauclair@ugent.be](mailto:Roxanne.Beauclair@ugent.be) (R. Beauclair).

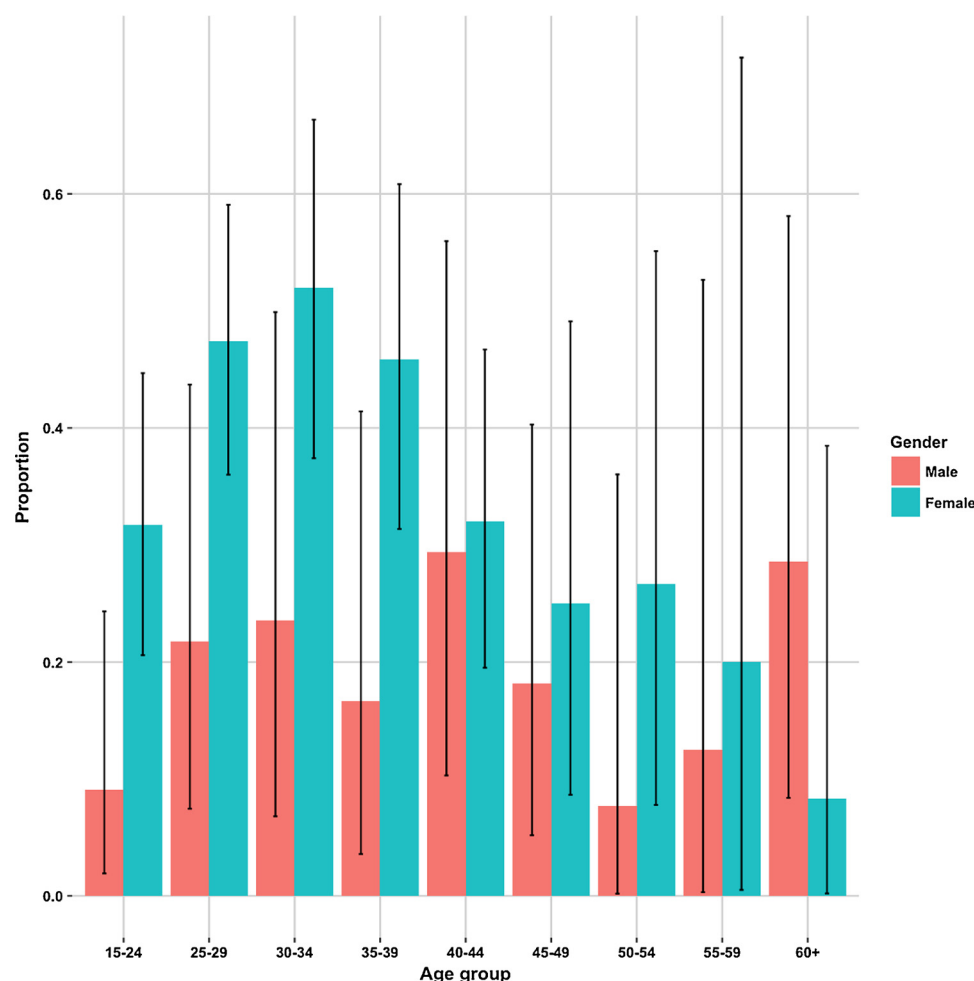


Fig. 1. HIV prevalence by sex and age in a randomly selected imputed dataset, with 95% exact confidence intervals.

et al., 2017). Recent findings from Southern Africa seem to suggest that specific partner age pairings play an important role in HIV transmission (Akullian et al., 2017; de Oliveira et al., 2016), and all relationships (not only primary or most recent relationships) are important for continuing the cycle of HIV transmission in a population (Beauclair et al., 2016; de Oliveira et al., 2016). Specifically, the phylogenetic study by de Oliveira et al., found that young women under 25 years old most likely acquired HIV infection from men 25–40 years old, and then as they aged and entered the 25–40-year-old age cohort they infected age-similar partners with HIV.

Complicating this is the role that variation in partner age may play in HIV transmission at the population level. A modelling study that investigated how age-mixing patterns – population-level trends for how people choose sexual partners with regards to age – affect HIV transmission, found that increased variance in partner ages for the whole population led to larger values of the basic reproduction number of HIV (d’Albis et al., 2012). Once the variance was large enough, the mean age difference in the population no longer influenced the transmission dynamics. The role of partner age variability in HIV transmission dynamics has been understudied in real-world populations (Beauclair et al., 2016).

Thus far in the age-mixing discourse most of the focus from observational studies has been on the individual-level risk posed by age differences in the most recent or primary relationship. Here, we are primarily interested in studying population-level attributes of age-mixing in real-world populations, and discussing how certain age-mixing patterns may create the conditions for HIV to be sustained in populations with high HIV prevalence. We also hypothesize that

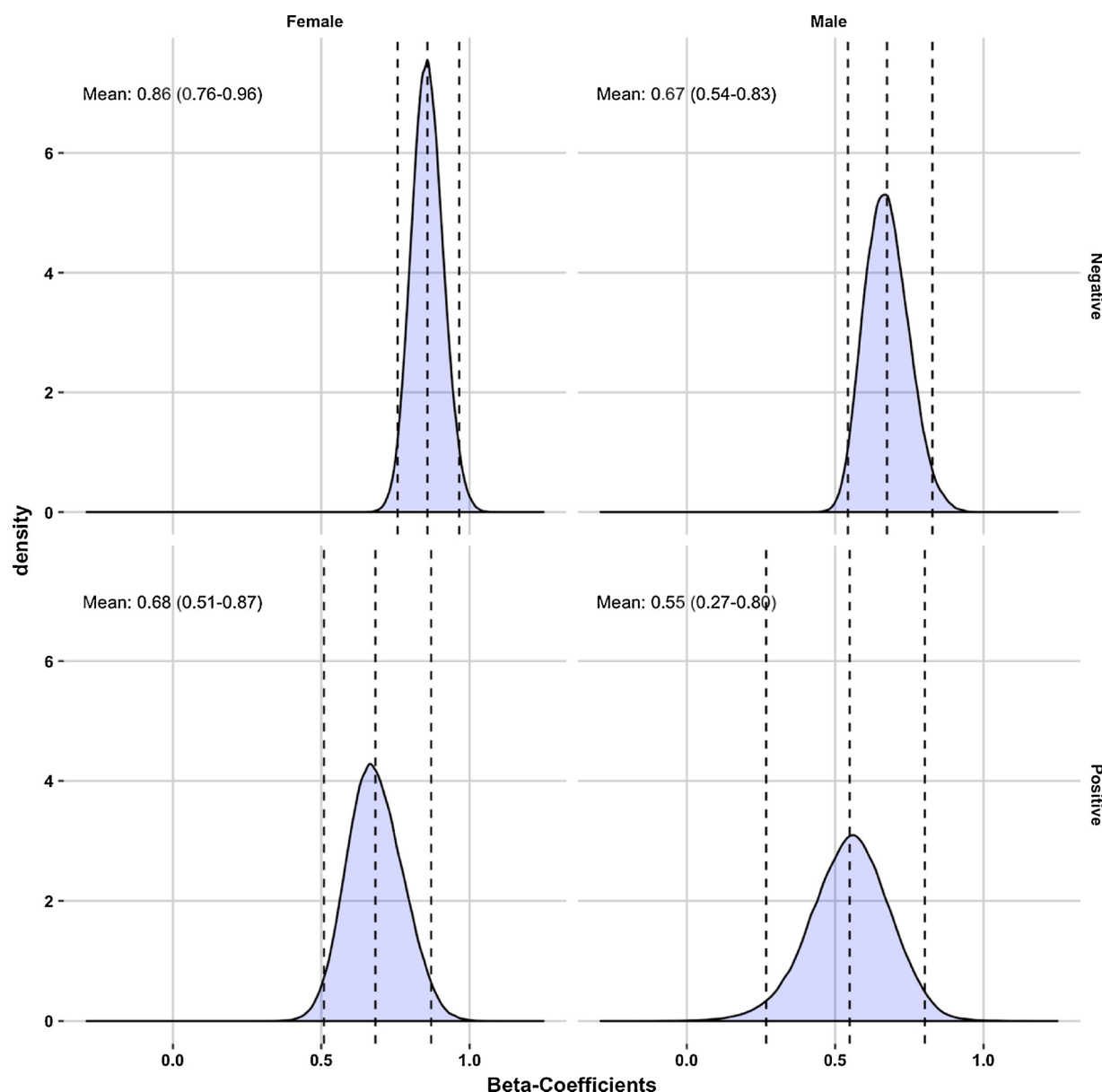
individuals with a large range of partner ages may allow HIV to be acquired from one age group and transmitted to another age group. It is important to note that the study of age-mixing among men has been largely absent from the literature. To this purpose we conducted an analysis of both men and women using data from a sexual relationship survey in Cape Town, South Africa. We aimed to describe the age-mixing patterns in HIV negative and HIV positive male and female participants and to investigate whether being HIV positive was correlated with having a larger range of partner ages during the year prior to the survey.

## 2. Methods

### 2.1. Data

The data for this analysis comes from a cross-sectional sexual behaviour survey we administered from June 2011 to February 2012 in three urban communities of Cape Town, South Africa that had a high HIV prevalence. The three study communities are designated “townships”: locations where “black” South Africans were allowed to live during Apartheid. These townships were characterized by high unemployment, low rates of graduation from secondary schools, as well as densely populated formal houses and shack dwellings (Department, 2012; Muzondo et al., 2004).

The survey employed Audio Computer-Assisted Self-Interviewing (ACASI) and collected in-depth sexual relationship histories of the participants for the year prior to the survey. Participants could report up to 15 partners.



**Fig. 2.** Distribution of model slopes, derived from the linear mixed-effects models. The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean of all the model slopes (beta-coefficients) represents the average increase in partner age for each year increase in participant age for that stratum.

Survey participants were randomly sampled from a roster of participants in the Zambia South Africa TB and AIDS Reduction (ZAMSTAR) prevalence survey, which took place in 2010, approximately one year before the sexual behaviour survey. In the ZAMSTAR prevalence survey the HIV status of participants was obtained via Abbott Determine HIV-1/2 screening tests (Ayles et al., 2008). These HIV test results were subsequently linked via a unique, anonymous ID to their sexual behaviour survey results. Full details of the Cape Town sexual behaviour survey study design, sampling procedures, as well as, contact and response rates of have been described elsewhere (Beauclair et al., 2015, 2013, Delva et al., 2013, 2011).

All together 878 participants took the sexual behaviour survey. We wanted to focus this analysis on the black, heterosexual and sexually active population, since this is one of the demographic groups most affected by HIV in South Africa. Therefore, we excluded participants whose stated sexual preferences were for the same or both genders ( $n = 56$ ), those who did not identify as black ( $n = 206$ ), and those not reporting any sexual partners in the previous year ( $n = 110$ ). This left

us with 506 participants reporting on 870 relationships.

## 2.2. Statistical analysis

Our data had missing observations for key participant-level variables: age ( $n = 17$ , 3.36%), sex ( $n = 13$ , 2.57%), and HIV status ( $n = 90$ , 17.79%). For partner age 7.47% of relationships ( $n = 65$ ) had missing values. A discussion of assumptions regarding the missingness mechanisms, as well as, comparisons of participant characteristics for key variables by missingness status can be found in the Supplementary Web Appendix 1. In order to account for the selection bias and loss of efficiency that otherwise would have been introduced by the missingness (Rubin, 1976) we imputed 100 datasets using multivariate imputation by chained equations with a random-forest algorithm (MICE-RF) (van Buuren et al., 2011; White et al., 2011). The random forest algorithm has been demonstrated to produce more accurate imputations when non-linear associations exist among variables used in analytic models (Penone et al., 2014; Stekhoven and Bühlmann, 2012). The

**Table 1**

Differences between the age-mixing pattern features, derived from the linear mixed-effects models. We calculated the differences between model features for HIV negative men (HNM) and HIV positive men (HPM), HIV negative women (HNW) and HIV positive women (HPW), HNM and HNW, as well as, HPM and HPW, for 100 imputed datasets, each bootstrapped 2000 times. We calculated 95% confidence intervals (CIs) using the percentile method.

Comparisons	$\beta$ -coefficient	Intercept	BSSD	WSSD
HNM vs. HNW	0.18 (0.00 — 0.35)	4.40 (0.78 — 8.36)	1.05 (-2.82 — 4.21)	1.01 (-1.11 — 3.28)
HPM vs. HPW	0.13 (-0.18 — 0.47)	2.10 (-7.23 — 10.50)	3.51 (-3.29 — 9.25)	1.47 (-3.62 — 6.17)
HPM vs. HNM	0.13 (-0.17 — 0.44)	4.43 (-4.01 — 13.85)	2.2 (-4.09 — 9.20)	2.03 (-2.02 — 6.54)
HPW vs. HNW	0.18 (-0.04 — 0.38)	2.13 (-1.87 — 6.10)	4.66 (1.58 — 7.86)	4.51 (0.90 — 7.79)

HNM, HIV Negative Men.

HNW, HIV Negative Women.

HPM, HIV Positive Men.

HPW, HIV Positive Women.

BSSD, Between-subject Standard Deviation.

WSSD, Within-subject Standard Deviation.

“mice” and “CALIBERrfimpute” packages in R were used for MICE-RF (Shah et al., 2014; van Buuren et al., 2011). A list of the variables we included in our imputation model can be found in Supplementary Web Appendix 2.

We wanted to describe the age-mixing pattern in four strata of participants: HIV positive men (HPM), HIV positive women (HPW), HIV negative men (HNM), and HIV negative women (HNW). To do this, we used linear mixed-effects models to regress the age of the partner on the age of the participant for each dataset. These models also accounted for the hierarchical nature of the data by containing a random intercept for the participant (Bates, 2010; Bolker et al., 2009; Gelman, 2007).

The average increase in partner age for each year increase in age of participant (model slope or  $\beta$ -coefficient), the average partner age for 15-year-olds (population-level intercept), the between-subject variability (between-subject standard deviation, BSSD), and the within-subject variability (residual, i.e. within-subject standard deviation, WSSD) were computed for each stratum using these models. We centred our model on 15 year olds because they represent the youngest age in the dataset and we wanted to understand how individuals who are only starting to become sexually active choose partners with respect to age. We aimed to compare these features of the age-mixing pattern for HNM vs. HPM, HNW vs. HPW, HNM vs HNW, and HPM vs HPW.

Since some features lack a standard calculation for the 95% confidence intervals (95% CIs) (e.g. MI-averaged WSSDs, or the differences between two WSSDs), we created 2000 participant-level bootstrap replicates for each imputed dataset ( $m = 100$ ), resulting in a total of 200,000 datasets (Davison and Hinkley, 1997; Efron, 1987). We applied the linear mixed-effects model for each population stratum of the 200,000 datasets and captured the age-mixing features. Then, we calculated the differences between those model features for the comparisons we were interested in and subsequently created 95% CIs using the percentile method.

Among the participants who reported more than one partner, we examined whether participant HIV status was associated with the range of partner ages. We created an indicator of this range, we call a “bridge width”. The bridge width is defined as the number of years difference between the maximum and the minimum partner age of the participant (Beauclair et al., 2016). First, we plotted the distribution of bridge widths. Then we treated bridge width as an over-dispersed count variable and regressed it on HIV status using negative binomial regression. We stratified the models by sex and adjusted for the age of the participants. We hypothesized that age was related to bridge width in a nonlinear way, and thus, we used a generalized additive model

smoothing for age (Hastie and Tibshirani, 1990; Wood, 2006). The exponentiated model coefficient for HIV status is referred to here as the expected bridge width ratio (EBWR). EBWRs and 95% CIs from the imputed datasets were pooled according to Rubin’s Rules (Rubin, 1987).

All analyses were conducted in the R statistical software (R Development Core Team, 2014). In this paper, we present the results from the analysis with imputed data. However, relationship characteristics from the original dataset, as well as the complete case analyses can be found in Supplementary Web Appendix 3.

### 3. Results

Of the relationships that did not have missing HIV status or gender data, 32.6% belonged to HNW ( $n = 219/671$ ), 35.7% to HNM ( $n = 240/671$ ), 24.7% to HPW ( $n = 166/671$ ), and 6.7% to HPM ( $n = 46/671$ ). Nearly 23% of relationships had either missing gender or HIV status (199/870).

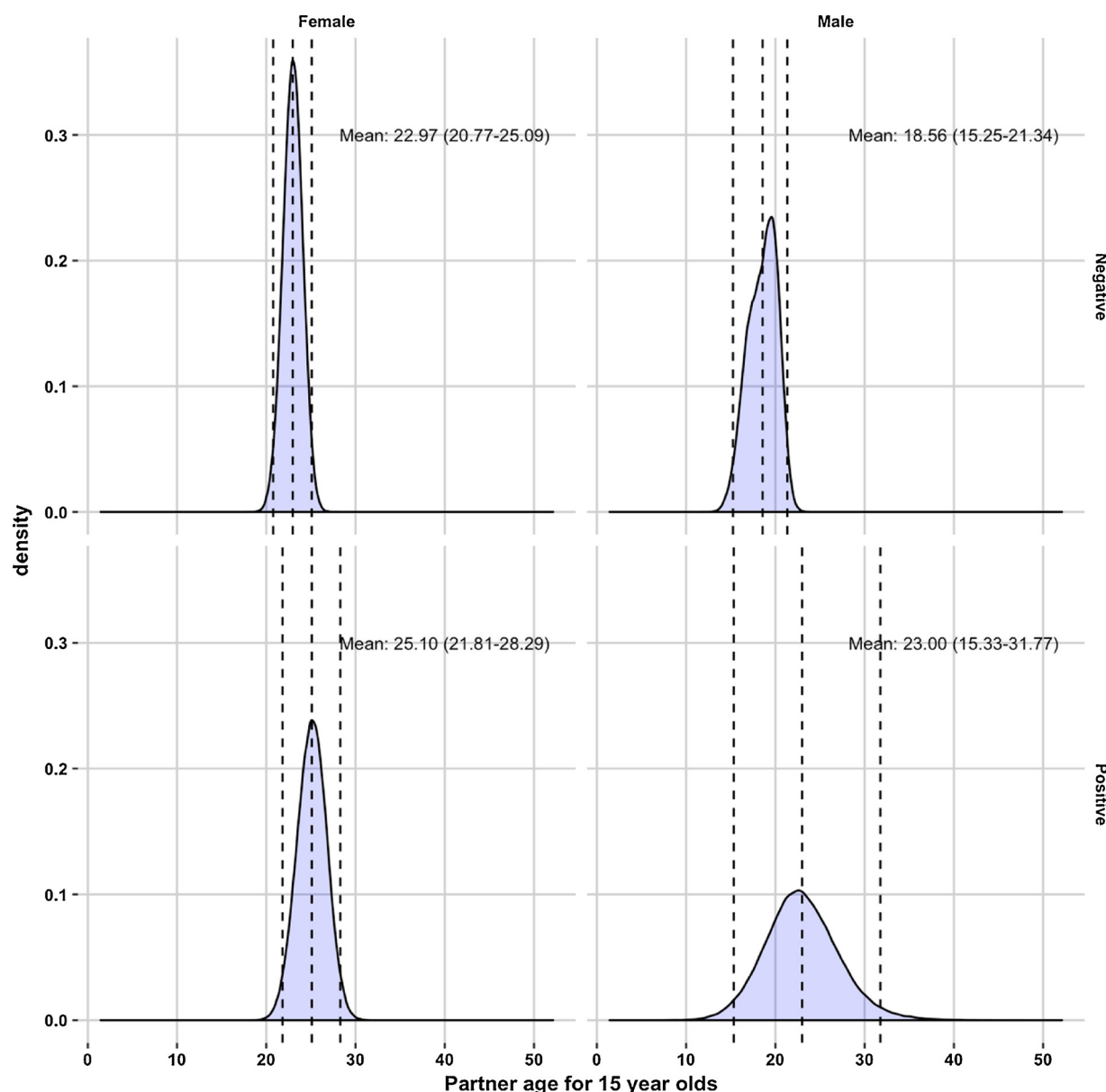
We randomly selected one of the 100 imputed data sets to visualize the HIV prevalence in our Cape Town study population (Fig. 1). In most age groups, women had a higher HIV prevalence than men. For women aged 15–24, HIV prevalence was nearly 3.5 times that of the men in that same age group (31.7% vs. 9.1%). This disparity was striking in 25–29-year-olds, with 47.4% of women infected and only 21.7% of men infected. HIV prevalence peaked in women 30–34 years old (52.0%) and in men who were 40–44 years old (29.4%).

Fig. 2 shows the distributions of model  $\beta$ -coefficients, which represent how fast the partner ages grow as participants’ ages increase. For all population strata, the average partner ages increased with each year increase in participant age. HPM had the smallest average model slope ( $\beta$ : 0.55; 95% CI: 0.27–0.8), while HNW had the largest ( $\beta$ : 0.86; 95% CI: 0.76–0.96) (Fig. 2). Most of the distributions seem to overlap to a great extent and in Table 1 we see there were only moderate differences in  $\beta$ -coefficients between the groups, with all but one of the 95% CIs around these differences overlapping with zero. HNW had steeper slopes than HNM (Difference: 0.18; 95% CI: 0–0.35). Also in Table 1 we see that for the HNM and HNW the model intercept difference was approximately 4 years (95% CI: 0.78–8.36), with 15-year-old men and women in those groups having partners who were 19 (95% CI: 15–21) and 23-years-old (95% CI: 21–25) on average, respectively (Fig. 3). If we take the mean model slopes and intercepts for each subpopulation the average partner age at 25 years old is predicted to be 32 years for HNW, 25 years for HNM, 32 years for HPW, and 28 years for HPM. The predicted partner ages for 40 year olds are: 44 years for HNW, 35 years for HNM, 42 years for HPW, and 37 years for HPM.

HNM had the largest BSSD (BSSD: 10.52; 95% CI: 6.97–12.95) while HPW had the lowest (BSSD: 4.81; 95% CI: 2.05–7.24) (Fig. 4). There appears to be a wide and bimodal distribution of BSSDs in HPM. In the complete case data, there were only 46 relationships in this subpopulation, with the minimum and maximum partner ages being 22 and 79, respectively. We believe those extreme values were over-sampled in some bootstrap replicates because of the small sample size, thus resulting in clusters of results with both small and large BSSDs. Table 1 shows that there was a relatively large and significant difference between the BSSD in HPW and HNW (Difference: 4.66; 95% CI: 1.58–7.86).

Fig. 5 shows that the WSSD was the largest for HPW (WSSD: 8.12; 95% CI: 4.81–11.04) and the smallest for HNW (WSSD: 3.62; 95% CI: 2.13–5.17), with the difference between the two groups being 4.51 (95% CI: 0.90–7.79) (Table 1). For the other group comparisons, the WSSDs appear to be similar and their distributions have a large degree of overlap.

In the complete study population 155 participants reported having had more than one relationship in the previous year (30.63%). Fig. 6 shows the distribution of bridge widths – range in partner ages – by population strata. In all subpopulations, the distributions were right-



**Fig. 3.** Distribution of model intercepts, derived from the linear mixed-effects models. The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean of all the model intercepts represents the average predicted partner age for 15-year-olds in that stratum.

skewed, with most of the bridge widths in all imputations spanning less than 20 years. In both HIV positive groups, there appears to be fatter tails, implying that those with HIV may have larger bridge widths, although there is more uncertainty in the imputations of HPM. The results of our individual-level negative-binomial models are presented in Figs. 7 and 8. Across all imputations, having an HIV positive status increased the expected bridge width of women by almost three times that of those who were HIV negative (EBWR: 2.76; 95% CI: 1.16–6.59). Among men there also seems to be an increased expectation of larger bridge widths in HIV positive participants compared to HIV negative participants, although when all of the estimates were pooled for the imputed data sets, the evidence did not appear to be as convincing as the evidence for women (EBWR: 1.93; 95% CI: 0.69–5.41).

Finally, in both men and women there appears to be a relationship between increasing participant age and larger bridge widths, although the uncertainty in the imputations is reflected by there being large variation in expected bridge widths at older ages. Indeed, in one randomly selected imputed dataset, there were only 22 participants (4.3%)

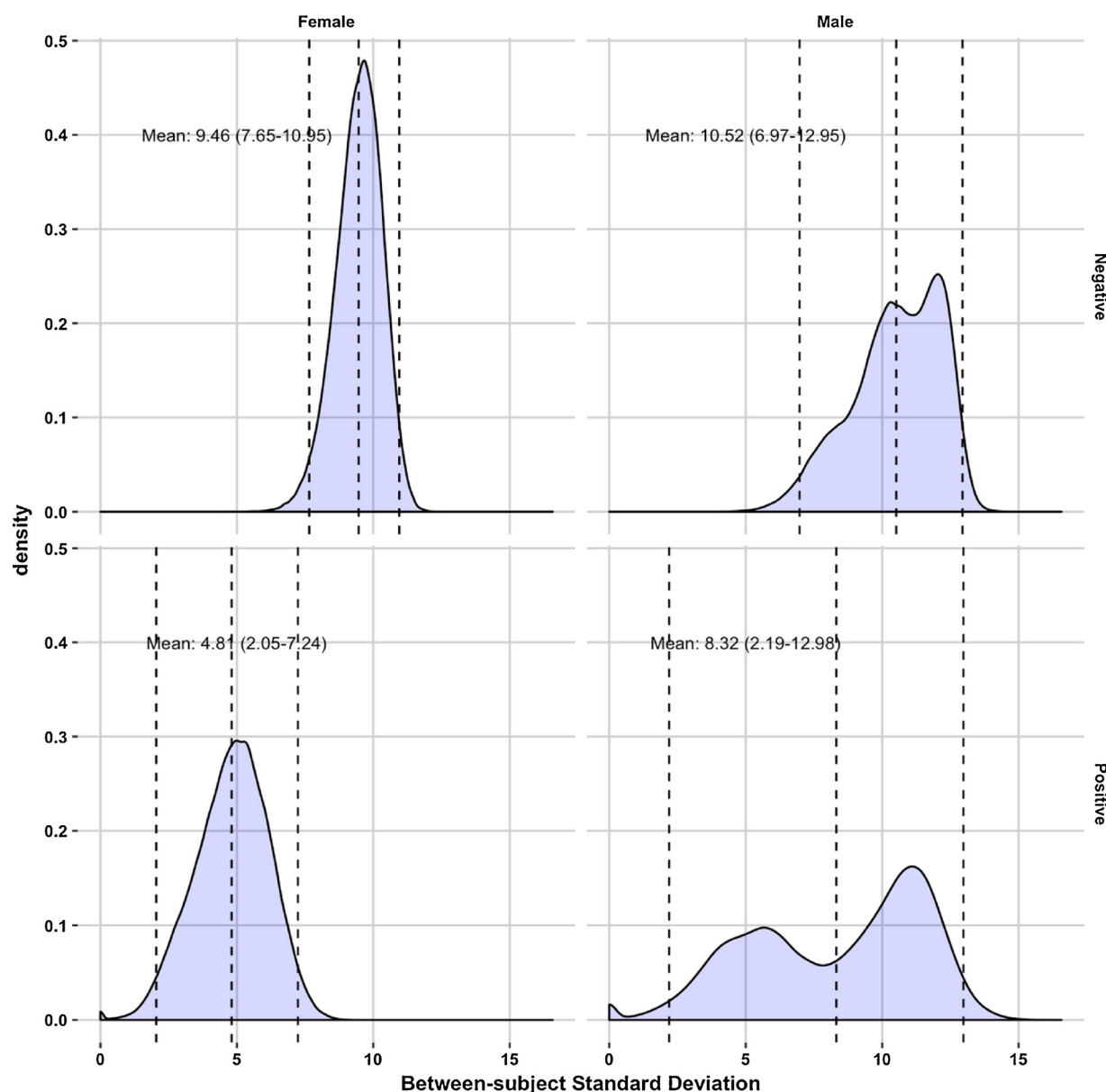
with ages between 60 and 70 years. For men, the increase in the expected partner age ranges were not as large as for women. For women, in most of the imputed data sets, the relationship between age and expected bridge widths was nonlinear, with the youngest and oldest participants having larger expected bridge widths, and those between 25 and 45 years old having smaller bridge widths.

#### 4. Discussion

Our study of heterosexual, black participants in urban Cape Town communities with high HIV prevalence sheds light on key features of the age-mixing patterns in HIV positive and negative men and women. In particular, there were some important findings with implications for the role women play in HIV transmission.

Perhaps the most novel insight from this study was that the within-subject variability of partner ages among all HPW was the largest out of all four population strata. Moreover, there was a detectable difference between their within-subject variability and that of the HNW. This





**Fig. 4.** Distribution of between-subject standard deviations (BSSD) of partner ages, derived from the linear mixed-effects models. The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean BSSD represents the average variability in partner age that can be seen among individuals in that stratum.

finding was also reinforced by our individual-level finding that HPW had nearly three times the expected range in partner ages compared to HNW. The HPM also had approximately two times the expected bridge width of HNM, though these results were not statistically significant. It is important to note that a large within-subject variability in partner ages does not necessarily mean that the women always have to have partners older than themselves. It is also possible that the HPW in our sample acquired HIV from age-similar partners. The large within-subject variability in partner ages in this group means that these women may have great potential to transmit HIV to men in other age cohorts, thereby allowing HIV to be sustained in a population (d'Albis et al., 2012). Interestingly, we witnessed the opposite effect for between-subject variability: the HPW had the smallest, and this was significantly different from the HNW. A larger fraction of HPW in our sample reported more than one partner compared to HNW. This may have resulted in the model being unable to distinguish which part of the overall variation in partner ages was due to the within component versus between component among HNW. Thus, the model might have attributed

too much variation to differences between participants among HNW. We believe more research is needed to explore the utility of BSSDs for understanding HIV transmission dynamics.

Another important finding is that for all population strata, 15-year-olds chose older partners, on average. However, in the HIV negative population, women chose significantly older partners than men did (23 years vs. 19 years). Critically, this demonstrates that in populations of young people, women are not the only ones choosing older partners. We also noted that in this study population the HIV prevalence among young women was many times greater than the HIV prevalence in age-similar men, indicating that if both young men and young women are choosing partners from their own age group the men will be more likely to choose an HIV positive partner, than the women. Most recent epidemiological studies examining the associations between age differences and HIV incidence have failed to include young men in their analyses, choosing to focus only on women in age-disparate relationships (Balkus et al., 2015; Harling et al., 2014; Schaefer et al., 2017; Street et al., 2015). However, there are a few studies that suggest that

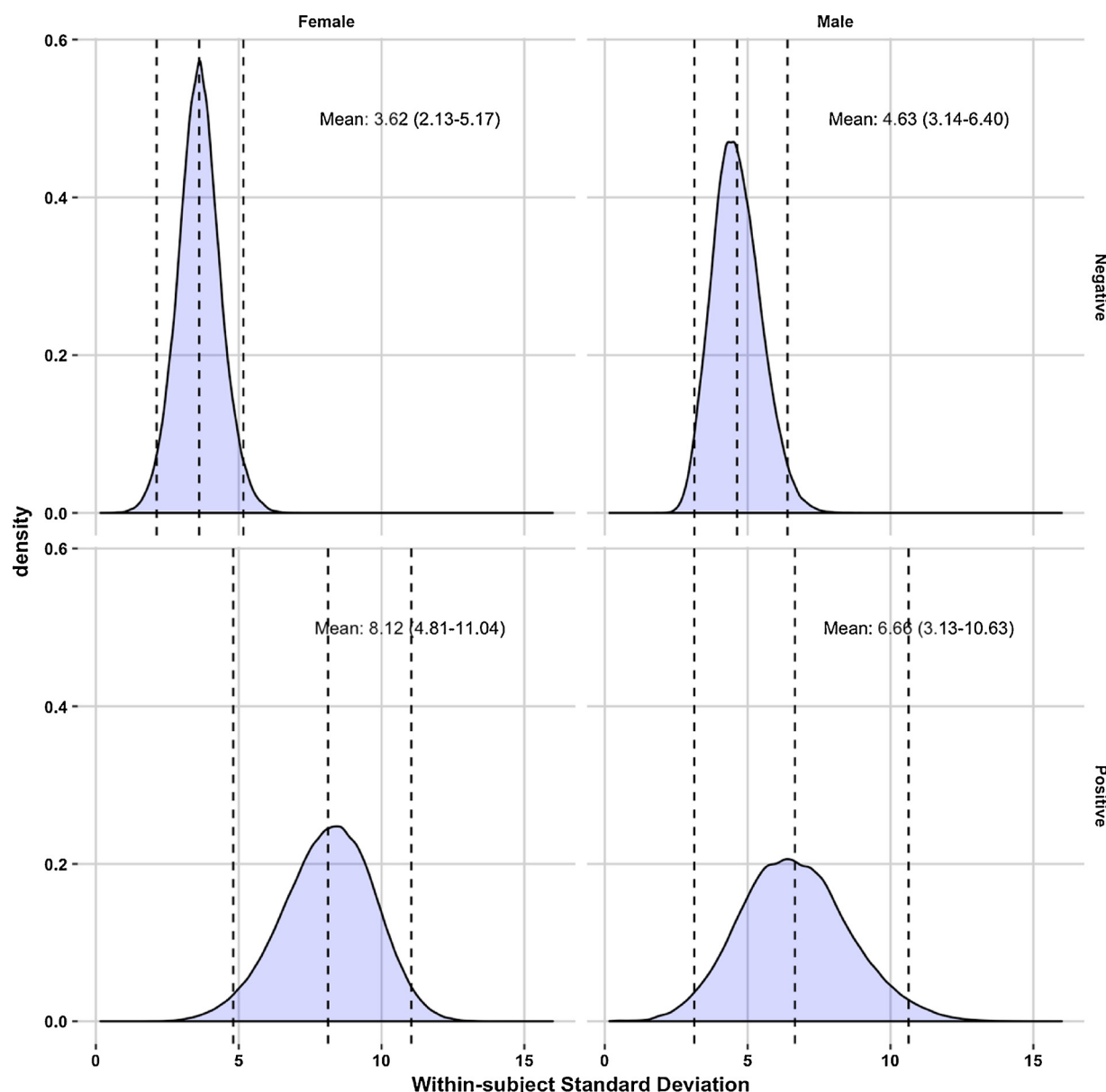


Fig. 5. Distribution of within-subject standard deviations (WSSD) of partner ages, derived from the linear mixed-effects models. The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean WSSD represents the average within-subject variability of partner age for individuals within each stratum.

men in relationships with older or age-similar women may have increased odds of having HIV or a curable STI (Beauclair et al., 2016; Gregson et al., 2002; O'Leary et al., 2015). We also note that men in urban communities are highly mobile, and they may be choosing female partners outside of the study communities. The women from those communities may not have the same HIV prevalence profiles as the women in our study community.

Among HIV-negative participants, there was a marginally significant difference between men and women for the effect of age on the average partner age choice. Indeed, we saw that young men in this group were younger than their female partners when they were 15, but by the time they were 40 years old they were on average 5 years older than their female partners, while women remained younger than their male partners at 40 years. This finding that as men get older, their age differences grow (older than the woman) is similar to other findings (Beauclair et al., 2016, 2012; Smith et al., 2014). While in other populations, women have been observed to have constant age differences as age increases (Beauclair et al., 2016, 2012; Becker et al., 2014), in

this study we observed, similar to Ott et al., that as women grew older, their age differences decreased (Ott et al., 2011). We did not detect significant differences between the HIV negative and the HIV positive participants of either gender, and thus we hypothesize that this particular feature of the age-mixing pattern may not play as important of a role in maintaining endemicity of HIV.

The results we present here are congruous with the findings from de Oliveira et al. who found in their phylogenetic analysis from KwaZulu-Natal, South Africa that HIV transmission pathways can be complex and involve cycles of transmission between different genders and age groups (de Oliveira et al., 2016). While our study was limited because we could not determine who infected whom, based upon evidence provided here, we hypothesize that young women below the age of 25, and older women above the age of 45, who have large bridge widths may be playing key roles in transmission. There have been many qualitative investigations about why young women are motivated to engage in age-disparate relationships (Beauclair and Delva, 2013; Hawkins et al., 2012; Leclerc-Madlala, 2008; Luke, 2003; Moore et al.,

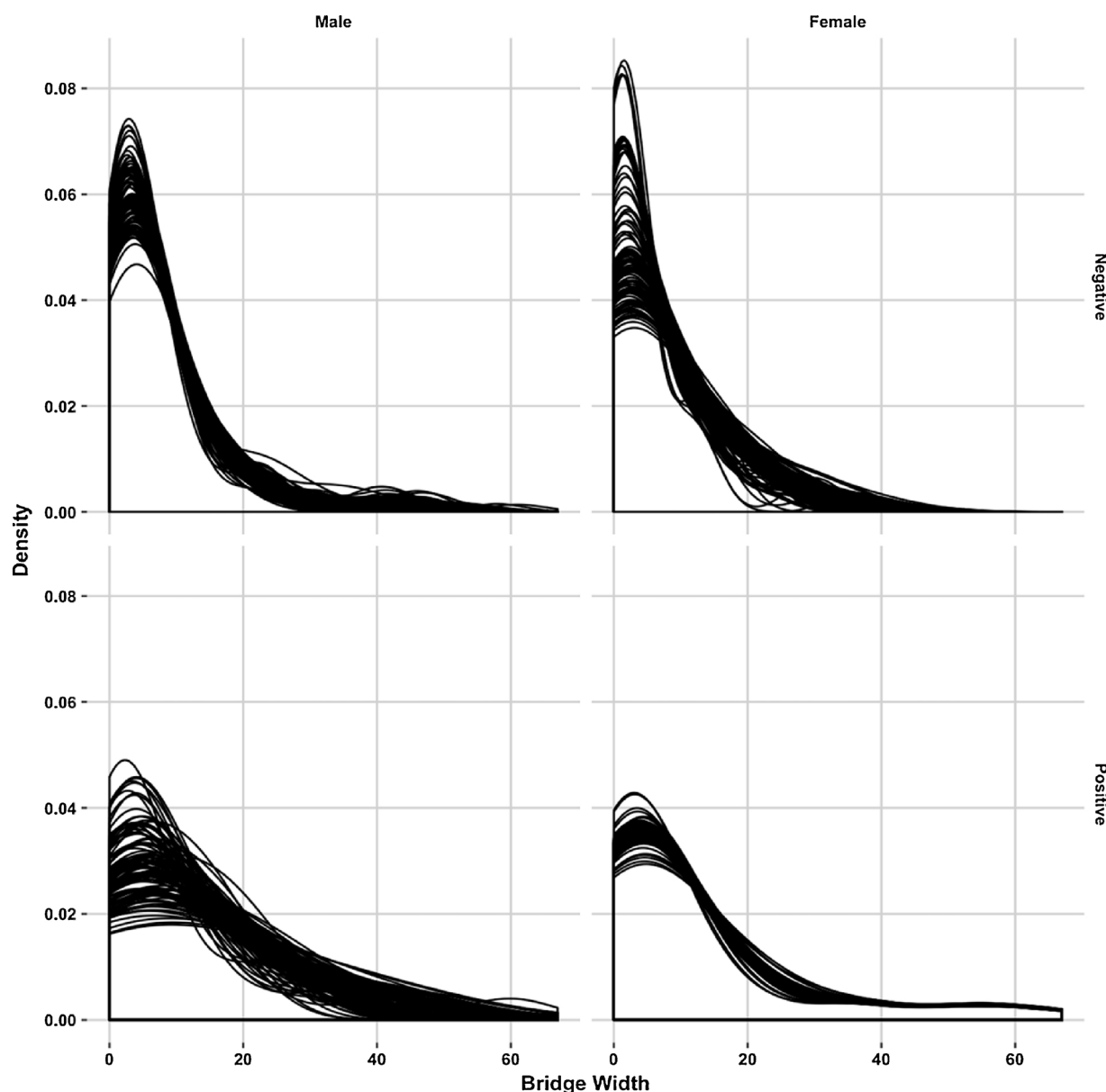


Fig. 6. Un-adjusted distribution of bridge widths among those reporting more than one partner for each imputed dataset, by sex and HIV status.

2007; Weiser et al., 2007). However, there is far less understanding about why older women have relationships with younger men, and vice versa (Kuate-Defo, 2008; Phaswana-Mafuya et al., 2014).

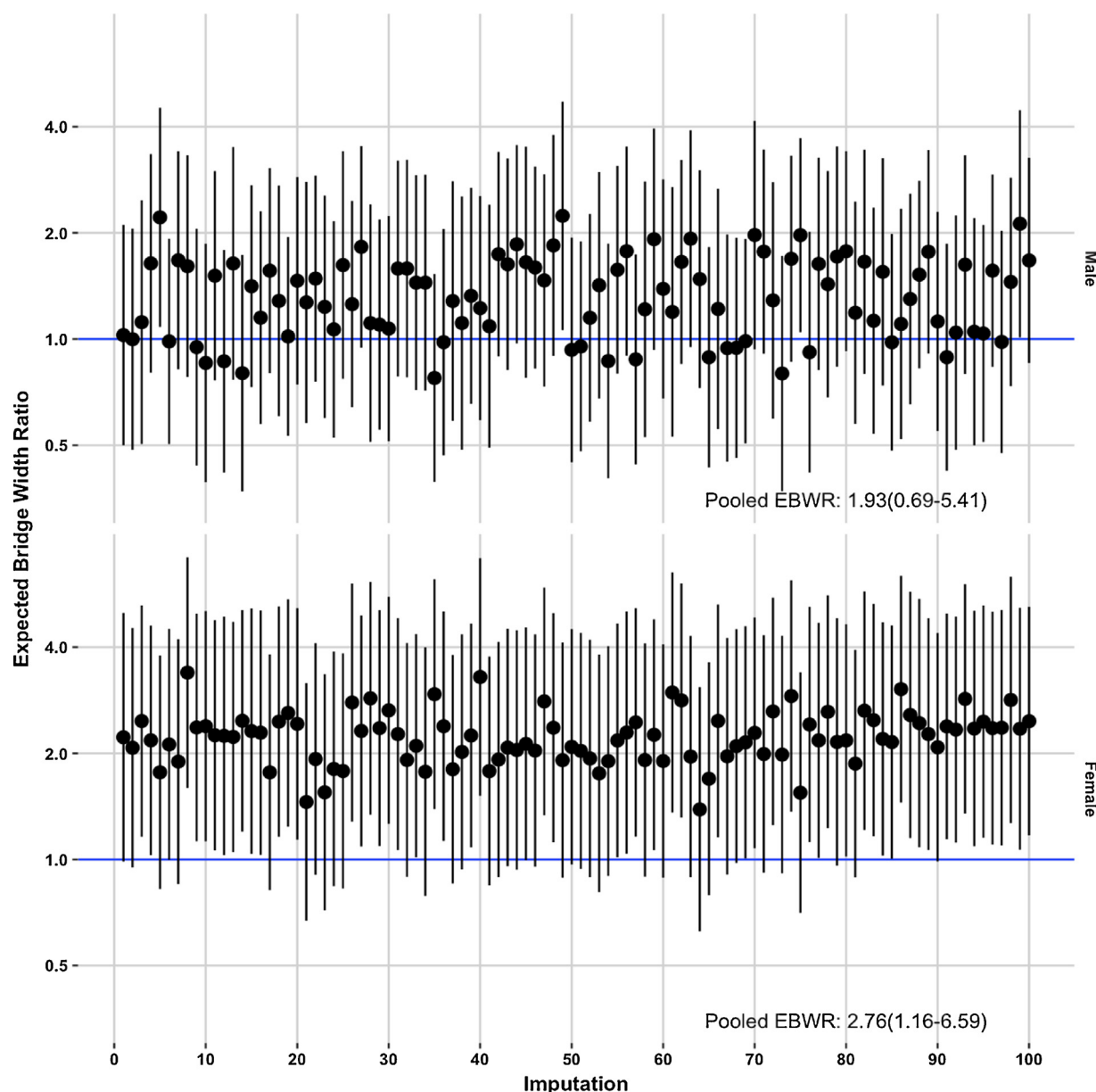
A key limitation of our study was that the sexual behaviour survey may have resulted in social desirability bias, through underreporting or over-reporting of sexual partners, in the cases of women and men, respectively. Additionally, one might expect that those with an HIV diagnosis would have been more likely to report age-similar partner ages, or refrain from answering questions about partner age. Fortunately, we were able to use techniques in the design of the survey, and in our analysis to counteract these forms of bias. We used ACASI methods to conduct the survey, which have been shown in other studies to reduce social desirability bias in different African settings (Dolezal et al., 2012; Hewett et al., 2004; Kissinger et al., 1999). Moreover, in a previously published paper using this data, we demonstrated how the ACASI interviewing modality likely resulted in more reporting of sexual partners and sensitive behaviours in comparison to other sexual behaviour surveys using face-to-face-interviewing (Beauclair et al., 2013). We also corrected for the bias that might have resulted from missing partner age data among HIV-positive participants, by using MICE-RF statistical

methods.

There is also a possibility that the HIV status variable may have had missing observations because the participants were HIV positive. In this scenario, MICE probably did not completely correct for that form of selection bias, and thus, a sensitivity analysis may be warranted. More research is necessary on how to conduct sensitivity analyses in the presence of MICE-RF imputations.

Additionally, we were restricted by the small total sample size, particularly of HIV positive participants. This may have resulted in a limited ability to detect differences between the HIV positive participants and other population strata. Unfortunately, it is often difficult to find large sexual behaviour datasets where participants have felt comfortable enough to report multiple partners in order to study these dynamics. A potential way forward is to conduct simulation studies, which do not have the sample size constraints of real-world data. Ultimately, it is important to note that these findings may only be generalizable to other urban, disadvantaged communities within Southern Africa that have similar HIV prevalence profiles among men and women of different age groups.





**Fig. 7.** Model coefficients for relationship between HIV and bridge width, for each of the 100 imputations. The Expected Bridge Width Ratio (EBWR) is the ratio of expected bridge widths for those who are HIV positive compared to those who are HIV negative. The EBWR is derived from the negative binomial model that also adjusts for the age of the participant. The EBWR that is annotated within the plots was pooled across imputations according to Rubin's Rules. The individual EBWRs that are plotted for each imputation, demonstrate the amount of variability seen in the imputations.

## 5. Conclusions

While age-mixing has garnered a lot of recent attention for the role it may play in HIV transmission, most of the studies have focused on the individual-level perspective of young women in age-disparate relationships. To our knowledge, this exploratory study is the first to try describe several real-world population-level features of age-mixing patterns in HIV negative and positive populations, and to see if they differ in systematic ways. These differences, or lack thereof, furnish insight into the complex transmission dynamics that occur between genders at different ages, and could possibly lead to sustained HIV epidemics. The hypotheses we generated about which features of age-mixing patterns are important for HIV epidemics, like within-subject variability, can be confirmed through simulation studies, calibrated to real-world populations such as the ones described here. These simulations could then be used to evaluate how different age-mixing pattern characteristics affect the individual-level and population-level

dynamics of HIV transmission over short and longer time scales.

## Funding sources

This work was supported by the Research Foundation - Flanders (FWO) [G091210N; G0B4314N; W002514N] and the Flemish Interuniversity Council (VLIR) [ZEIN2010PR375].

## Acknowledgements

We would like to thank the fieldworkers who helped to carry out this study. We also acknowledge assistance from Nulda Beyers and Rory Dunbar from the Desmond Tutu TB Foundation at Stellenbosch University. They were instrumental with helping to design the study, providing us with a sampling frame and subsequently linking the ZAMSTAR HIV results to our sexual behaviour survey.

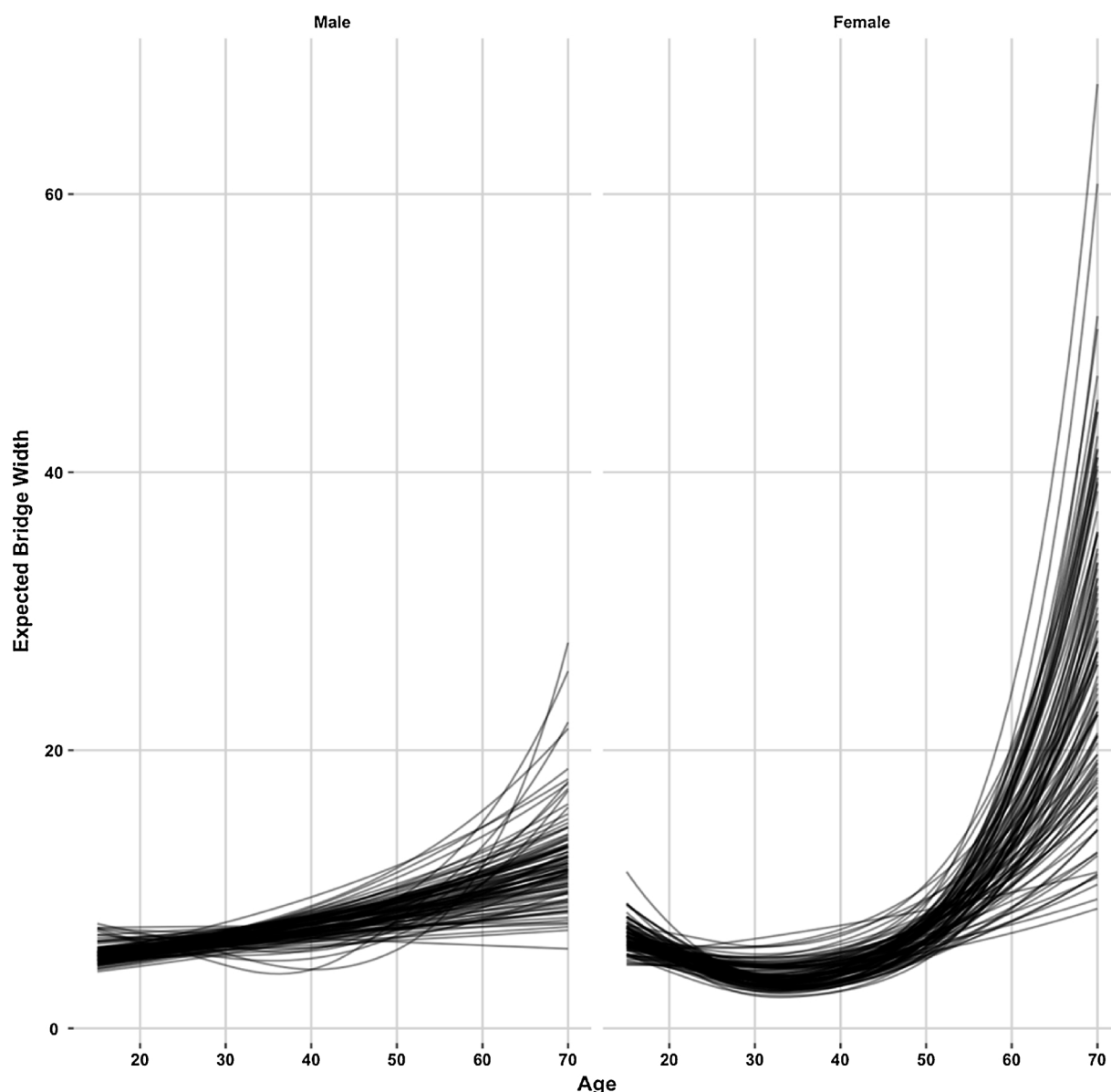


Fig. 8. Expected bridge widths for different values of age (smooth term), by gender. The expected bridge widths are predicted from the negative binomial model that also adjusts for the HIV status of participants. There is a line for each of the 100 models resulting from the imputations.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.epidem.2018.05.006>.

## References

- Akullian, A., Bershteyn, A., Klein, D., Vandormael, A., Barnighausen, T., Tanser, F., 2017. Sexual partnership age-pairings and risk of HIV acquisition in rural South Africa. *AIDS* 31, 1755–1764. <http://dx.doi.org/10.1097/QAD.0000000000001553>.
- Anderson, R.M., May, R.M., Ng, T.W., Rowley, J.T., 1992. Age-dependent choice of sexual partners and the transmission dynamics of HIV in sub-Saharan Africa. *Philos. Trans. Biol. Sci.* 336, 135–155.
- Ayles, H.M., Sismanidis, C., Beyers, N., Hayes, R.J., Godfrey-Faussett, P., 2008. ZAMSTAR, the Zambia South Africa TB and HIV reduction study: design of a 2 x 2 factorial community randomized trial. *Trials* 9.
- Balkus, J.E., Nair, G., Montgomery, E.T., Mishra, A., Palanee-Phillips, T., Ramjee, G., Panchia, R., Selepe, P., Richardson, B.A., Chirenje, Z.M., Marrazzo, J.M., 2015. Age-disparate partnerships and risk of HIV-1 acquisition among South African women participating in the VOICE trial. *J. Acquir. Immune Defic. Syndr.* 70, 212–217. <http://dx.doi.org/10.1097/QAI.0000000000000715>.
- Bates, D.M., 2010. *lme4: Mixed-Effects Modelling With R*. Springer.
- Beauclair, R., Delva, W., 2013. Is younger really safer? A qualitative study of perceived risks and benefits of age-disparate relationships among women in Cape Town, South Africa. *PLoS One* 8.
- Beauclair, R., Hellingranger, S., Hens, N., Delva, W., 2016. Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi. *Sci. Rep.* 6, 36121. <http://dx.doi.org/10.1038/srep36121>.
- Beauclair, R., Hens, N., Delva, W., 2015. Concurrent partnerships in Cape Town, South Africa: race and sex differences in prevalence and duration of overlap. *J. Int. AIDS Soc.* 18, 19372. <http://dx.doi.org/10.7448/IAS.18.1.19372>.
- Beauclair, R., Kassanjee, R., Temmerman, M., Welte, A., Delva, W., 2012. Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa. *Eur. J. Contracept. Reprod. Heal. Care* 17, 30–39. <http://dx.doi.org/10.3109/13625187.2011.644841>.
- Beauclair, R., Meng, F., Deprez, N., Temmerman, M., Welte, A., Hens, N., Delva, W., 2013. Evaluating audio computer assisted self-interviews in Urban South African communities: evidence for Good suitability and reduced social desirability bias of a Cross-sectional survey on sexual behaviour. *BMC Med. Res. Methodol.* 13, 11.
- Becker, M., Vignoles, V.L., Owe, E., Easterbrook, M.J., Brown, R., Smith, P.B., Bond, M.H., Regalia, C., Manzi, C., Brambilla, M., Aldhafri, S., Gonzalez, R., Carrasco, D., Paz Cadena, M., Lay, S., Schweiger Gallo, I., Torres, A., Camino, L., Ozgen, E., Guner, U.E., Yamakoglu, N., Silveira Lemos, F.C., Trujillo, E.V., Balanta, P., Macapagal, M.E., Cristina Ferreira, M., Herman, G., de Sauvage, I., Bourguignon, D., Wang, Q., Fulop, M., Harb, C., Chybicka, A., Mekonnen, K.H., Martin, M., Nizharadze, G., Gavreliuc, A., Buitendach, J., Valk, A., Koller, S.H., 2014. Cultural bases for self-evaluation: seeing oneself positively in different cultural contexts. *Pers. Soc. Psychol. Bull.* 40, 657–675. <http://dx.doi.org/10.1177/0146167214522836>.

- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* 24.
- Chapman, R., White, R.G., Shafer, L.A., Pettifor, A., Mugurungi, O., Ross, D., Pascoe, S., Cowan, F.M., Grosskurth, H., Buve, A., Hayes, R.J., 2010. Do behavioural differences help to explain variations in HIV prevalence in adolescents in sub-Saharan Africa? *Trop. Med. Int. Heal.* 15, 554–566.
- d'Albis, H., Augeraud-Veron, E., Djemai, E., Ducrot, A., 2012. The dispersion of age differences between partners and the asymptotic dynamics of the HIV epidemic. *J. Biol. Dyn.* 6, 695–717. <http://dx.doi.org/10.1080/17513758.2012.688146>.
- Davison, A., Hinkley, D., 1997. *Bootstrap Methods and Their Application*, vol. 1. Cambridge University Press.
- de Oliveira, T., Kharsany, A.B.M., Gräff, T., Cawood, C., Khanyile, D., Grobler, A., Puren, A., Madurai, S., Baxter, C., Karim, Q.A., Karim, S.S.A., 2016. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. *Lancet HIV* 3018, 1–10. [http://dx.doi.org/10.1016/S2352-3018\(16\)30186-2](http://dx.doi.org/10.1016/S2352-3018(16)30186-2).
- Delva, W., Beauclair, R., Welte, A., Vansteelandt, S., Hens, N., Aerts, M., du Toit, E., Beyers, N., Temmerman, M., 2011. Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health* 11, 616. <http://dx.doi.org/10.1186/1471-2458-11-616>.
- Delva, W., Meng, F., Beauclair, R., Deprez, N., Temmerman, M., Welte, A., Hens, N., 2013. Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa. *J. Int. AIDS Soc.* 16, 18034. <http://dx.doi.org/10.7448/IAS.16.1.18034>.
- Department, S.D.I. and G., 2012. *City of Cape Town – 2011 Census – Cape Town*. Cape Town.
- Dolezal, C., Marhefka, S.L., Santamaria, E.K., Leu, C.S., Brackis-Cott, E., Mellins, C.A., 2012. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch. Sex Behav.* 41, 401–410. <http://dx.doi.org/10.1007/s10508-011-9769-6>.
- Efron, B., 1987. Better bootstrap confidence intervals. *J. Am. Stat. Assoc.* 82, 171–200.
- Gelman, A., 2007. *Data Analysis Using Regression and multilevel/hierarchical Models*. Cambridge University Press, New York.
- Gouws, E., Stanecki, K.A., Lyster, R., Ghys, P.D., 2008. The epidemiology of HIV infection among young people aged 15–24 years in southern Africa. *Aids* 22, S5–S16 doi: 10.1097/01.aids.0000341773.86500.9dr00002030-200812004-00002 [pii].
- Gregson, S., Nyamukapa, C.A., Garnett, G.P., Mason, P.R., Zhuwau, T., Carael, M., Chandiwana, S.K., Anderson, R.M., 2002. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet* 359, 1896–1903 doi: S0140-6736(02)08780-9 [pii] 10.1016/S0140-6736(02)08780-9.
- Hallett, T.B., Gregson, S., Lewis, J.J., Lopman, B.A., Garnett, G.P., 2007. Behaviour change in generalised HIV epidemics: impact of reducing cross-generational sex and delaying age at sexual debut. *Sex Transm. Infect.* 83 (Suppl. 1), i50–54 doi:sti.2006.023606 [pii] 10.1136/sti.2006.023606.
- Harling, G., Newell, M.L., Tanser, F., Kawachi, I., Subramanian, S.V., Barnighausen, T., 2014. Do age-disparate relationships Drive HIV incidence in Young women? Evidence from a population cohort in rural KwaZulu-Natal, South Africa. *J. Acquir. Immune Defic. Syndr.* 66, 443–451. <http://dx.doi.org/10.1097/QAI.0000000000000198>.
- Hastie, T.J., Tibshirani, R., 1990. *Generalised Additive Models*. Chapman and Hall/CRC, Boca Raton.
- Hawkins, K., Price, N., Mussa, F., 2012. Milking the cow: Young women's construction of identity and risk in age-disparate transactional sexual relationships in Maputo, Mozambique. *Glob. Public Heal. Int. J. Res. Policy Pract.* 4, 169–182.
- Hewett, P.C., Mensch, B.S., Erulkar, A.S., 2004. Consistency in the reporting of sexual behavior by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm. Infect.* 80 (Suppl. 2), ii43–ii48. <http://dx.doi.org/10.1136/sti.2004.013250>.
- Kissinger, P., Rice, J., Farley, T., Trim, S., Jewitt, K., Margavio, V., Martin, D.H., 1999. Application of computer-assisted interviews to sexual behavior research. *Am. J. Epidemiol.* 149, 950–954.
- Kuate-Defo, B., 2008. Young people's relationships with sugar daddies and young people's relationships sugar mummies : what do we know and what do we need to know? *Afr. J. Reprod. Health* 8, 13–37.
- Leclerc-Madlala, S., 2008. Age-disparate and intergenerational sex in Southern Africa: the dynamics of hypervulnerability. *AIDS* 22 (Suppl. 4), S17–S25 doi:10.1097/01.aids.0000341774.86500.5300002030-200812004-00003 [pii].
- Luke, N., 2003. Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Stud. Fam. Plann.* 34, 67–86.
- Maughan-brown, B., Evans, M., George, G., 2016. Sexual behaviour of men and women within age-disparate partnerships in South Africa : implications for Young women 's HIV risk. *PLoS One* 11–16. <http://dx.doi.org/10.1371/journal.pone.0159162>.
- Maughan-Brown, B., Kenyon, C., Lurie, M.N., 2014. Partner age differences and concurrency in South Africa: implications for HIV-infection risk among Young women. *AIDS Behav.* <http://dx.doi.org/10.1007/s10461-014-0828-6>.
- Moore, A.M., Biddlecom, A.E., Zulu, E.M., 2007. Prevalence and meanings of exchange of money or gifts for sex in unmarried adolescent sexual relationships in sub-Saharan Africa. *Afr. J. Reprod. Health* 11, 44–61.
- Morrison-Beedy, D., Xia, Y., Passmore, D., 2013. Sexual risk factors for partner age discordance in adolescent girls and their male partners. *J. Clin. Nurs.* 22, 3289–3299. <http://dx.doi.org/10.1111/jocn.12408>.
- Muzondo, I.F., Barry, M., Dewar, D., Whittall, J., Africa, S., Muzondo, I.F., Barry, M., 2004. Athens, Greece. Land Conflicts in Informal Settlements: Wallacedene in Cape Town, South Africa, FIG Working Week 2004.
- O'Leary, A.E., Jemmott III, J.B., Jemmott, L.S., Teitelman, A., Heeren, G.A., Ngwane, Z., Icard, L.D., Lewis, D.A., 2015. Associations between psychosocial factors and incidence of sexually transmitted disease among South African adolescents. *Sex Transm. Dis.* 42, 135–139.
- Ott, M.Q., Barnighausen, T., Tanser, F., Lurie, M., Newell, M.L., 2011. Age-gaps in sexual partnerships: seeing beyond 'sugar daddies'. *AIDS* 25, 861–869.
- Penone, C., Davidson, A.D., Shoemaker, K.T., DiMarco, M., Rondinini, C., Brooks, T.M., Young, B.E., Graham, C.H., Costa, G.C., 2014. Imputation of missing data in life-history trait datasets: which approach performs best? *Methods Ecol. Evol.* 5, 961–970.
- Pettifor, A.E., Rees, H.V., Kleinschmidt, I., Steffenson, A.E., MacPhail, C., Hlongwa-Madikizela, L., Vermaak, K., Padian, N.S., 2005. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *AIDS* 19, 1525–1534 doi: 00002030-200509230-00012 [pii].
- Phaswana-Mafuya, N., Shisana, O., Davids, A., Tabane, C., Mbelle, M., Matseke, G., Banyini, M., Kekana, Q., 2014. Perceptions of sugar mommy practices in South Africa. *J. Psychol. Afr.* (South Sahara, Caribb., Afro-Lat. Am. 24, 257–263. <http://dx.doi.org/10.1080/14330237.2014.906086>.
- R Development Core Team, 2014. *R: A Language and Environment for Statistical Computing*.
- Ritchwood, T.D., Hughes, J.P., Jennings, L., MacPhail, C., Williamson, B., Selin, A., Kahn, K., Gómez-Olivé, F.X., Pettifor, A., 2016. Characteristics of age-discordant partnerships associated with HIV risk among Young South African women (HPTN 068). *JAIDS J. Acquir. Immune Defic. Syndr.* 72, 423–429. <http://dx.doi.org/10.1097/QAI.0000000000000988>.
- Rubin, D., 1987. *Multiple Imputation for Nonresponse in Surveys*. Wiley, New York.
- Rubin, D.B., 1976. Inference and missing data. *Biometrika* 63 581–592. <http://dx.doi.org/10.1093/biomet/63.3.581>.
- Schaefer, R., Gregson, S., Eaton, J.W., Mugurungi, O., Rhead, R., Takarua, A., Maswera, R., Nyamukapa, C., 2017. Age-disparate relationships and HIV incidence in adolescent girls and young women: evidence from a general-population cohort in Zimbabwe. *AIDS* 31 (10), 1461. <http://dx.doi.org/10.1097/QAD.0000000000001506>.
- Shah, A.D., Bartlett, J.D., Carpenter, J., Nicholas, O., Hemingway, H., 2014. Comparison of random Forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. *Am. J. Epidemiol.* 179, 764–774.
- Shisana, O., Rehle, T., Simbayi, L., Zuma, K., Jooste, S., Zungu, N., Labadarios, D., Onoya, D., 2014. *South African National HIV Prevalence, Incidence and Behaviour Survey, 2012*. Cape Town.
- Smith, J., Nyamukapa, C., Gregson, S., Lewis, J., Magutshwa, S., Schumacher, C., Mushati, P., Hallett, T., Garnett, G., 2014. The distribution of sex acts and condom use within partnerships in a rural sub-Saharan African population. *PLoS One* 9, e88378. <http://dx.doi.org/10.1371/journal.pone.0088378>.
- Stekhoven, D.J., Buhlmann, P., 2012. MissForest - nonparametric missing value imputation for mixed-type data. *Bioinformatics* 28, 112–118.
- Street, R.A., Reddy, T., Ramjee, G., 2015. The generational effect on age disparate partnerships and the risk for human immunodeficiency virus and sexually transmitted infections acquisition. *Int. J. STD AIDS* 27 (9), 746–752. <http://dx.doi.org/10.1177/0956462415592325>.
- van Buuren, S., Groothuis-Oudshoorn, K., Buuren, S., Groothuis-Oudshoorn, K., 2011. MICE: multivariate imputation by chained equations in R. *J. Stat. Softw.* 45, 1–67.
- Weiser, S.D., Leiter, K., Bangsberg, D.R., Butler, L.M., Percy-de Korte, F., Hlanze, Z., Phaladze, N., Iacopino, V., Heisler, M., 2007. Food insufficiency is associated with high-risk sexual behavior among women in Botswana and Swaziland. *PLoS Med.* 4, 1589–1597. <http://dx.doi.org/10.1371/journal.pmed.0040260>. discussion 1598.
- White, I.R., Royston, P., Wood, A.M., 2011. Multiple imputation using chained equations: issues and guidance for practice. *Stat. Med.* 30, 377–399. <http://dx.doi.org/10.1002/sim.4067>.
- Wood, S., 2006. *Generalized Additive Models: An Introduction with R*. Taylor & Francis.